(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) Publication number: 0 357 370 A2

(12) # EUROPEAN PATENT APPLICATION

(21) Application number: 89308718.9

(22) Date of filing: 29.08.89

(51) Int. Cl.5: G 06 F 15/38

(30) Priority: 02.09.88 JP 220715/88

(43) Date of publication of application: 07.03.90 Bulletin 90/10

(84) Designated Contracting States: DE GB

(71) Applicant: SHARP KABUSHIKI KAISHA
22-22 Nagaike-cho Abeno-ku
Osaka 545 (JP)

(72) Inventor: Fukumochi, Yoji
2-7-12, Higashishigigaoka Sango-cho
Ikoma-gun Nara-ken (JP)

Kugimiya, Shuzo
44-401 1, Ukyo 2-chome
Nara-shi Nara-ken (JP)

Namamura, Ichiko
Takanohara-Urban C-408, 2-2, Kabutodai Kizu-cho
Soraku-gun Kyoto-fu (JP)

Hirai, Tokuyuki
492, Minosho-cho
Yamatokoriyama-shi Nara-ken (JP)

The other inventor has agreed to waive his entitlement to designation

(74) Representative: Brown, Kenneth Richard et al
R.G.C. Jenkins & Co. 26 Caxton Street
London SW1H 0RJ (GB)

(54) **Computer assisted language translating machine.**

(57) A computer assisted language translating machine comprising a morpheme analyzer for separating a source language sentence, transmitted from a communication appliance, into morphemes; a syntax analyzer for analyzing the syntax of a string of the morphemes; a converter for converting a syntactic structure of the input sentence into a syntactic structure of the target language; a sentence generator for generating a translated sentence according to the syntactic structure of the target language; a first storage unit for storing character data forming the source language sentence; a character data writing device for sequentially writing the character data; a character data reading device for reading the character data from the first storage unit after a predetermined time subsequent to the writing of the character data into the first storage unit; a second storage unit for storing the character data read out from the first storage unit; a discriminator for determining whether or not the character data read out by the character data reading device represents the end of one sentence; and a translation starting means operable, when the discriminator indicates that the character data read out represents the end of one sentence, to initiate a translation with respect to the character data representative of such one sentence and stored in the second storage unit.

EP 0 357 370 A2

Description

BACKGROUND OF THE INVENTION

1. Field of the Invention

The present invention relates to a computer assisted language translating machine and, more particularly, to the translating machine capable of accomplishing a real-time translation of an original text received from a communication appliance such as, for example, a telex machine.

2. Description of the Prior Art

Where an original text in a source language, for example, English, received by a communication appliance is translated into a translated version in a target language, for example, Japanese, with the use of the computer assisted language translating machine, an operator of the translating machine has to manually input a message, received by the communication appliance and representing the original text, through a keyboard arrangement into the translating machine. Alternatively, the translating machine can be connected with the communication appliance so that the translating machine can receive the transmitted message directly from the communication appliance and, in such case, the transmitted message has to be temporarily stored in a system storage file. When the message stored in the system storage file is to be translated, the operator has to undergo a time-consuming manipulation to call the message from the storage file and then to enter it in a translation module in the translating machine.

Either the entry of the original text into the translating machine or the withdrawal of the received message from the system storage file for the entry thereof into the translating machine requires the intervention of manual labors. Accordingly, the prior art method is ineffective to accomplish a real-time translation of the message in one source language into a translated version in a target language different from the source language.

SUMMARY OF THE INVENTION

The present invention has been developed with a view to substantially solving the above described disadvantages and has for its essential object to provide an improved computer assisted language translating machine capable of accomplishing a real-time translation of a message in a source language, received from a communication appliance, into a translated version in a target language.

According to the present invention, there is provided a computer assisted language translating machine which comprises a morpheme analyzer for separating a source language sentence, transmitted from a communication appliance, into morphemes to provide information including parts of speech of the morphemes with the use of at least a fixed main dictionary; a syntax analyzer for analyzing the syntax of a string of the morphemes, which have been separated by the morpheme analyzer, with the use of a dictionary and grammatical rules; a converter for converting a syntactic structure of the input sentence, analyzed by the syntax analyzer, into a syntactic structure of the target language; and a sentence generator for generating a translated sentence according to the syntactic structure of the target language converted by the converter.

To accomplish the above described object of the present invention, the computer assisted language translating machine also comprises a first storage unit for storing character data forming the source language sentence; a character data writing means for sequentially writing the character data in the order in which the character data are inputted from the communication appliance; a character data reading means for reading the character data from the first storage unit in the order, in which the character data are sequentially stored in the first storage unit, after a predetermined time subsequent to the writing of the character data into the first storage unit; a second storage unit for storing the character data read out by the character data reading means from the first storage unit; a discriminating means for determining whether or not the character data read out by the character data reading means represents the end of one sentence; and a translation starting means operable, when the discriminating means indicates that the character data read out represents the end of one sentence, to initiate a translation with respect to the character data representative of such one sentence and stored in the second storage unit.

According to the present invention, when the character data representing the sentence received from the communication appliance is inputted, the character data are sequentially written into the first storage unit in the order in which the character data have been inputted. Subsequently, after the passage of the predetermined time subsequent to the writing of the character data into the first storage unit performed by the character writing means, the character data written in the first storage unit are read out by the character data reading means in the order, in which they have been written, for the storage into the second storage means. At this time, the discriminating means determines whether or not the character data read out by the character data reading means represent the end of one sentence. Should the discriminating means determine that the character data representative of the end of one sentence have been read out, the translation starting means initiates the translation with respect to the character data representative of such one sentence and stored in the second storage means.

Then, when the source language sentence is inputted in the morpheme analyzer, the source language sentence is separated into a strings of morphemes to provide information including the parts of speech. The syntactic structure of the morpheme strings is then analyzed by the syntax analyzer with the use of the dictionary and the grammatical rules and is subsequently converted by

the converter into a syntactic structure of the target language. The syntactic structure of the target language is utilized by the sentence generator for generating a translated sentence according to such syntactic structure of the target language.

Accordingly, when the character data of the source language sentence received from the communication appliance is inputted, a translation thereof into the target language is automatically executed sentence to sentence.

BRIEF DESCRIPTION OF THE DRAWINGS

This and other objects and features of the present invention will become apparent from the following description taken in conjunction with a preferred embodiment thereof with reference to the accompanying drawings in which like parts are designated by like reference numerals and in which:

Fig. 1 is a block diagram of a preferred embodiment of a computer assisted language translating machine according to the present invention;

Fig. 2 is an operating structure diagram of a translation module according to the preferred embodiment;

Fig. 3 is a diagram showing the contents stored in a buffer A;

Fig. 4 is a flowchart showing the sequence of operation for storing character data; and

Fig. 5 is a flowchart showing the sequence of reading and translating data inputted from a communication appliance.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENT

Referring to Fig. 1, a block diagram of a computer assisted language translating machine according to the present invention is shown. The translating machine comprises a main CPU (central processing unit) 1, a main memory 2, a CRT 3 (cathode ray tube) 3, a keyboard 4, a translation module 5, a memory device 6 storing dictionaries, grammatical rules and tree structure conversion rules for translation, and a printer for providing a visible output representative of a result of translation performed by the computer assisted language translating machine.

When a source language text is inputted to the translation module 5, the translation module 5 translates and outputs a target language counterpart. Specifically, the source language text entered from keyboard 4 is transferred to the translation module 5 under the control of the main CPU 1. The translation module 5 makes use of grammatical rules and tree structure conversion rules stored in the memory 6 to translate the inputted source language text into the target language counterpart as will be described later. The result of this translation is outputted to the printer to provide a translated version of the source language text. The translated version may also be displayed on the CRT 3.

Fig. 2 illustrate a specific structure of the translation module 5 used in the practice of the preferred embodiment of the present invention.

When the source language sentence is inputted in the morpheme analyzer 11, the morpheme analyzer 11 makes use of the dictionary stored in the memory 6 to separate the source language sentence into a strings of morphemes (word string) to provide information including the parts of speech of each of words forming the source language sentence, information on the grammar and translations of each word and also to analyze tense, person and number. The syntactic structure of the morpheme strings is then analyzed by the syntax analyzer 12 which makes use of the grammatical rules and the dictionary to determine a syntax analysis tree structure determinative of the relationship between the individual words forming the sentence.

The converter 13 makes use of the tree structure conversion rules stored in the memory 6 to convert the structure of the syntax analysis tree for the source language sentence to the structure of the target language sentence. The sentence generator 14 applies the Japanese postpositional article and auxiliary verbs appropriate to the Japanese translation, the resultant translation being outputted from the translation module 5 to the printer 7.

Fig. 3 illustrates the structure of a buffer A for the storage of character date comprising the source language text transmitted from the communication appliance such as, for example a telex machine. The buffer A is adapted to store character data representative of each character at a location specified by a writing pointer which is incremented at the time of writing the character data. When reading the written character data out from the buffer A, the character representative of each character is read out from the location specified by a reading pointer which is incremented at the time of reading the character data.

Fig. 4 illustrates a flowchart showing the sequence of operation for storing character data and Fig. 5 illustrates a flowchart showing the sequence of reading and translating data inputted from a communication appliance.

Hereinafter, a process for the real-time translation according to the preferred embodiment of the present invention will now be described with particular reference to Figs. 3 to 5. Referring first to Fig. 4, at step S1, the writing pointer indicating the location in the buffer A at which the character data is to be stored is zeroed for initializing the character data storage buffer A. Then, at step S2, and when the character data are transmitted from the communication appliance such as the telex machine, a decision is made to determine if one character data so inputted represents character data indicative of a RETURN data (line-begin command). If a result of decision at step S2 indicates that the character data represents the RETURN data, the program flow proceeds to step S5, but if it does not, the program flow proceeds to step S3.

At step S3, the character data so inputted is written in the buffer A at the location specified by the writing pointer and, thereafter, the writing pointer is incremented at step S4.

At step S5, a decision is made to determine if the character storage process should be continued. Where it should be continued, the program flow proceeds to step S6, but if it should not, the

character data storage process is terminated.

At step S6, a decision is made to determine if the writing pointer is greater than MAX (indicative of the last location for the storage of the character data in the buffer A). In the event that the writing pointer is greater than Max, it means that the buffer A is full of the character data and, therefore, the program flow return to step S1 to reset the writing pointer to zero. On the other hand, if the writing pointer is not greater than Max, the program flow returns to step S2 to continue the character data storage process for the next succeeding character data.

On the other hand, when the storage of the character date representative of the source language text into the buffer A is initiated in the manner as hereinabove described, the reading of the character data so stored in the buffer A takes place after a predetermined time subsequent to the initiation of the storage of the character data representative of the source language text. Hereinafter, the sequence of reading and translating data inputted from the communication appliance will be described with reference to Fig. 5.

At step S11, the reading pointer is set to zero, followed by step S12 at which the character data corresponding to one character is read out from the location of the buffer A specified by the reading pointer. At step S13, a decision is made to determine if the character data corresponding to one character read from the buffer A represents a character data indicative of a period (.). Should it represents the period, the program flow proceeds to step S17, but should it not represent the period, the program flow proceeds to step S14.

At step S14, the character data corresponding to one character read out at step S12 is stored in a buffer B, not shown, for the storage of a source language from which translation should be made. Then, at step S15, a decision is made to determine if the reading pointer is equal to MAX. If it is equal to MAX, the program flow return to step S11, but if it is not, the program flow proceeds to step S16.

At step S16, the reading pointer is incremented, followed by step S12.

At step S17, the character data representative of the period read out is stored in the buffer B. Then, at step S18, a decision is made to determine if the reading pointer is equal to MAX. If it is equal to MAX, the program flow proceeds to step S20.

At step S19, the reading pointer is set to zero, followed by step S21.

At step S20, the reading pointer is incremented, followed by step S21.

At step S21, the character data corresponding to one character is read out from the location of the buffer A specified by the reading pointer, followed by step S22 at which the character data so read out is examined as to whether or not it represents a space. If it represents the space, the program flow proceeds to step S24, but if it does not, the program flow proceeds to step S23. At step S23, the character date corresponding to one character read out in the manner as hereinabove described is stored in the buffer B and, thereafter, step S27 takes place.

Since the character data representative of the period (.) has been read out at step S12 and the character data representative of the space has also been read out at step S21, all of the character data corresponding to one sentence are determined as read out from the buffer A and subsequently stored in the buffer B and, accordingly, the translation of the character data so stored in the buffer B is initiated at step S24.

With the translation process so initiated, the character data corresponding to one sentence referred to above are inputted to the translation module 5 under the control of the main CPU 1 for the translation thereof into the target language.

At subsequent step 25, the resultant translation given at step S24 is outputted to the printer 7 to provide a printed output of the resultant translation of one sentence, thereby completing one cycle of translation process.

At step S27, a decision is made to determine if the reading pointer is equal to MAX. If it is equal to MAX, the program flow return to step S11 to set the reading pointer to zero, but if it is not, the program flow proceeds to step S28 at which the reading pointer is incremented and the program flow subsequently return to step S12 to read the character data corresponding to the next succeeding one sentence out from the location of the buffer A specified by the reading pointer.

Thus, according to the present invention, when the character data representative of the sentence in the source language is inputted from the communication appliance such as the telex machine, the character data are temporarily stored in the buffer A thereby to avoid any possible leak of portion of the character data transmitted direct from the communication appliance. Simultaneously therewith, the input of the character data from the communication appliance provides a cue with which the reading process for reading the character data stored in the buffer A is automatically started after the predetermined time. Then, the end of one message is automatically recognized and a translation program stored in the translation module is started to execute the translation without requiring the intervention of the operator.

According to the present invention, it is possible to accomplish the real-time translation once the source language message received by the communication appliance is inputted to the translating machine.

Thus, from the foregoing description, it is clear that, when the character data representing the sentence received from the communication appliance is inputted, the character data are sequentially written into the first storage unit. The character data written in the first storage unit are subsequently read out by the character data reading means. At this time, the discriminating means determines whether or not the character data read out by the character data reading means represent the end of one sentence and, in the event that the discriminating means determine that the character data representative of the end of one sentence have been read out, the translation starting means initiates the

translation with respect to the character data representative of such one sentence and stored in the second storage means. Accordingly, when the character data of the source language sentence received from the communication appliance is inputted, a translation thereof into the target language is automatically executed sentence to sentence.

Although the present invention has been fully described in connection with the preferred embodiment thereof with reference to the accompanying drawings, it is to be noted that various changes and modifications are apparent to those skilled in the art. For example, although in describing the preferred embodiment of the present invention, reference has been made to the example wherein the English sentence is translated into Japanese, the concept of the present invention can be equally applicable to a translation of Japanese, or any other source language, into English or any other target language linguistically different from the source language.

The invention being thus described, it will be obvious that the same may be varied in many ways. Such variations are not to be regarded as a departure from the scope of the invention. There are described above novel features which the skilled man will appreciate give rise to advantages. These are each independent aspects of the invention to be covered by the present application, irrespective of whether or not they are included within the scope of the following claims.

Claims

1. A computer assisted language translating machine comprising:
a morpheme analyzer for separating a source language sentence, transmitted from a communication appliance, into morphemes to provide information including parts of speech of the morphemes with the use of at least a fixed main dictionary;
a syntax analyzer for analyzing the syntax of a string of the morphemes, which have been separated by the morpheme analyzer, with the use of a dictionary and grammatical rules;
a converter for converting a syntactic structure of the input sentence, analyzed by the syntax analyzer, into a syntactic structure of the target language;
a sentence generator for generating a translated sentence according to the syntactic structure of the target language converted by the converter;
a first storage unit for storing character data forming the source language sentence;
a character data writing means for sequentially writing the character data in the order in which the character data are inputted from the communication appliance;
a character data reading means for reading the character data from the first storage unit in the order, in which the character data are sequentially stored in the first storage unit, after a predetermined time subsequent to the writing of the character data into the first storage unit;
a second storage unit for storing the character data read out by the character data reading means from the first storage unit;
a discriminating means for determining whether or not the character data read out by the character data reading means represents the end of one sentence; and
a translation starting means operable, when the discriminating means indicates that the character data read out represents the end of one sentence, to initiate a translation with respect to the character data representative of such one sentence and stored in the second storage unit.

2. Computer assisted language translating apparatus adapted to provide a real-time translation of character data received in a first language from a communications device into character data in a second language.

3. Computer assisted language translating apparatus comprising:
first memory means;
writing means for sequentially writing character data from a communications device into the first memory means;
second memory means;
means for transferring said character data from the first memory means to the second memory means;
means for detecting the end of each sentence of character data as it is transferred to the second memory means; and
translating means operable in response to said detecting means to translate a sentence of the character data stored in the second memory means once the sentence end is detected.

4. A computer assisted language translating machine comprising: a morpheme analyzer for separating a source language sentence, transmitted from a communication appliance, into morphemes; a syntax analyzer for analyzing the syntax of a string of the morphemes; a converter for converting a syntactic structure of the input sentence into a syntactic structure of the target language; a sentence generator for generating a translated sentence according to the syntactic structure of the target language; a first storage unit for storing character data forming the source language sentence; a character data writing device for sequentially writing the character data; a character data reading device for reading the character data from the first storage unit after a predetermined time subsequent to the writing of the character data into the first storage unit; a second storage unit for storing the character data read out from the first storage unit; a discriminator for determining whether or not the character data read out by the character data reading device represents the end of one sentence; and a translation starting means operable, when the discriminator indicates that the character data read out represents the end of one sentence, to initiate a translation with respect to the charac-

ter data representative of such one sentence and stored in the second storage unit.

5. A computer assisted language translating machine comprising a morpheme analyzer for separating a source language sentence, transmitted from a communication appliance, into morphemes to provide information including parts of speech of the morphemes with the use of at least a fixed main dictionary; a syntax analyzer for analyzing the syntax of a string of the morphemes, which have been separated by the morpheme analyzer, with the use of a dictionary and grammatical rules; a converter for converting a syntactic structure of the input sentence, analyzed by the syntax analyzer, into a syntactic structure of the target language; and a sentence generator for generating a translated sentence according to the syntactic structure of the target language converted by the converter.

Fig. 1

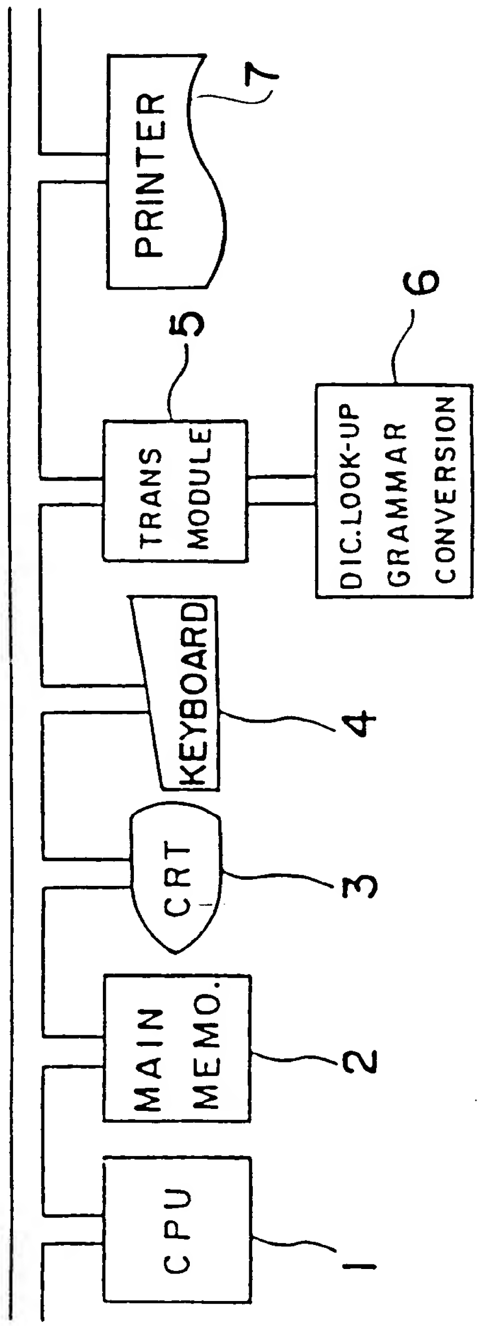

CPU
1
MAIN MEMO.
2
CRT
3
KEYBOARD
4
TRANS MODULE
5
DIC.LOOK-UP GRAMMAR CONVERSION
6
PRINTER
7

Fig. 2

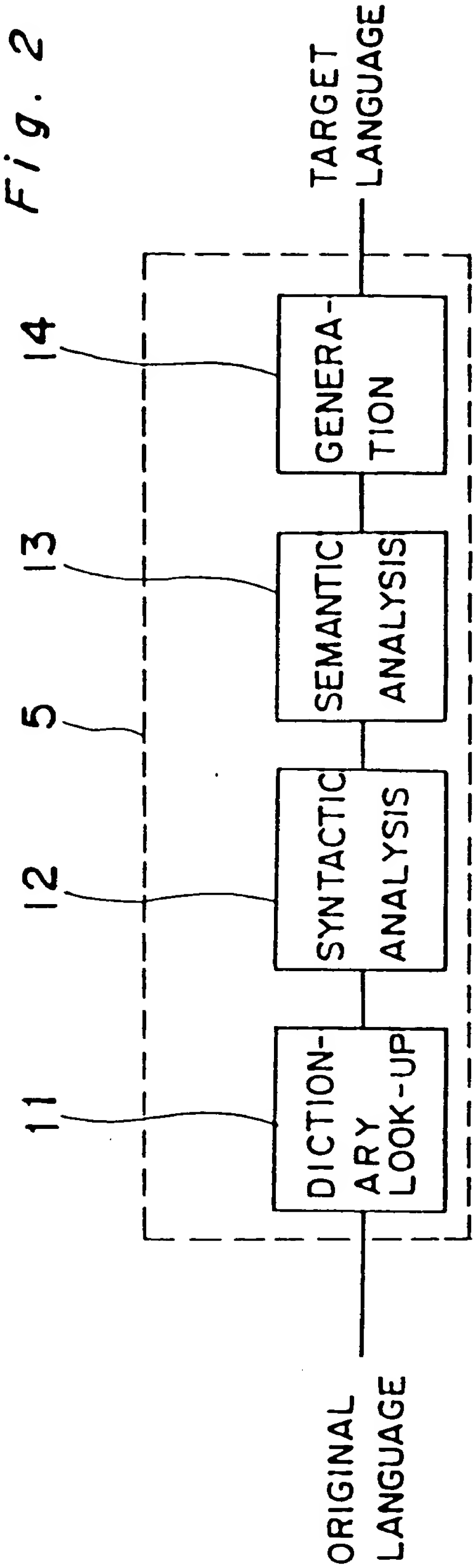

ORIGINAL LANGUAGE
DICTION-ARY LOOK-UP
11
SYNTACTIC ANALYSIS
12
5
SEMANTIC ANALYSIS
13
GENERA-TION
14
TARGET LANGUAGE

Fig. 3

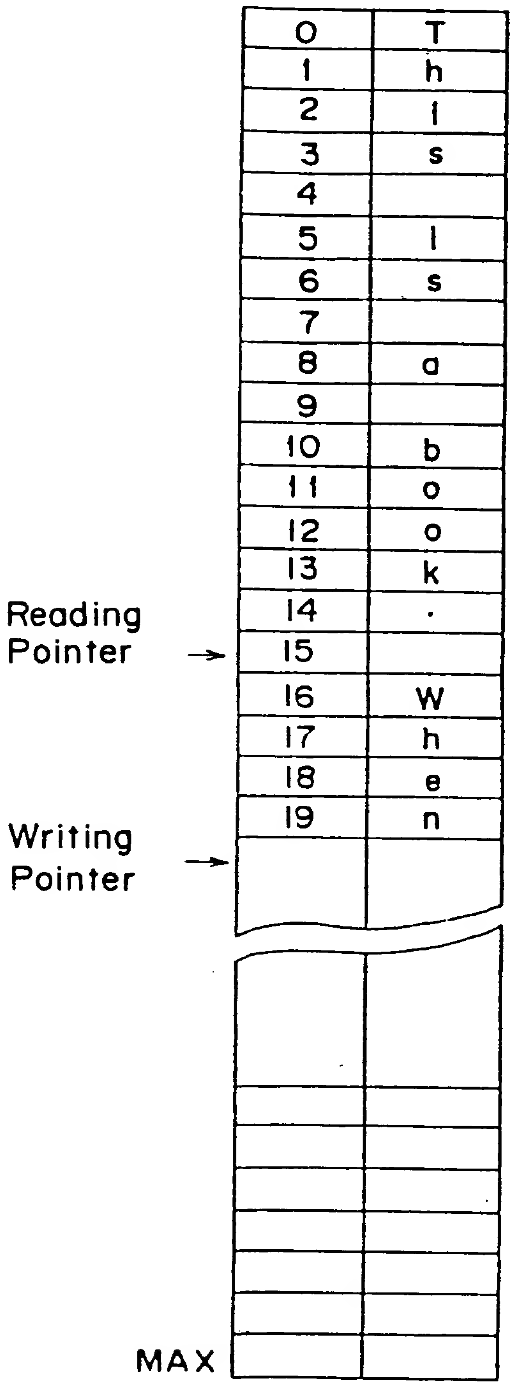
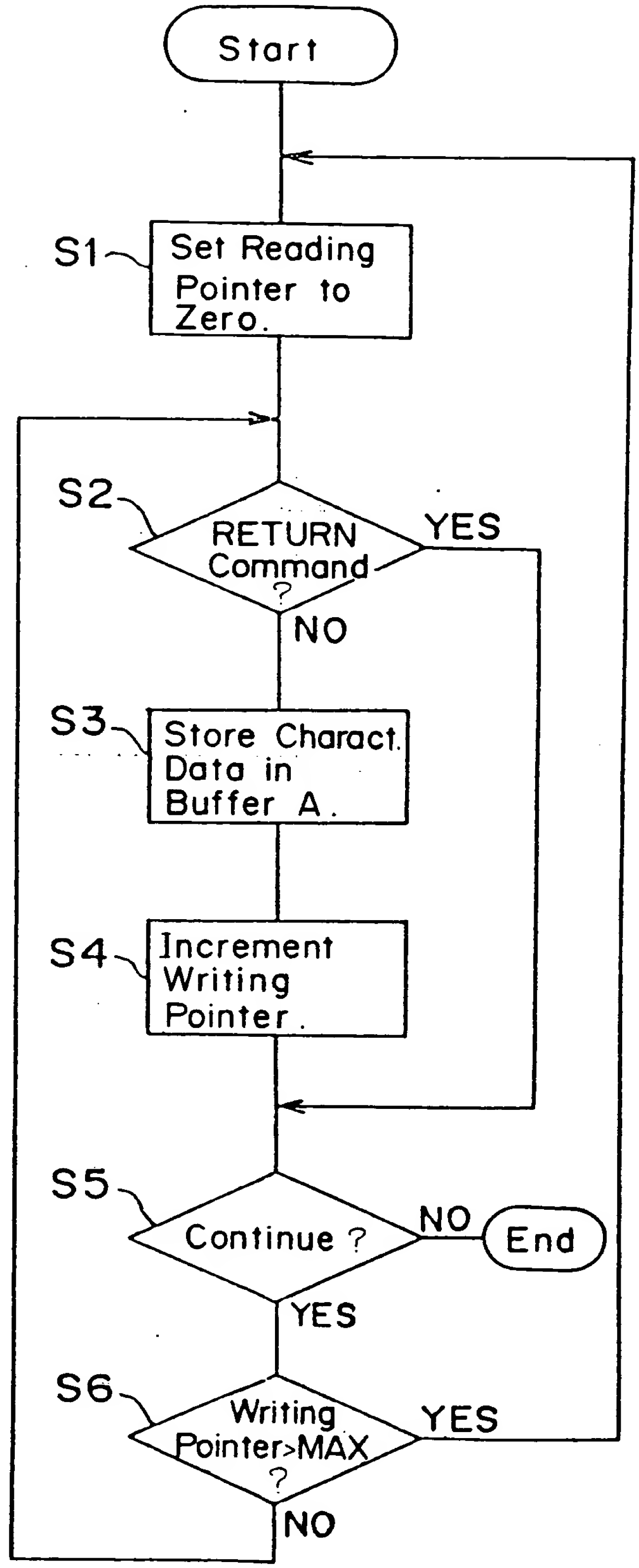

0 T
1 h
2 i
3 s
4
5 i
6 s
7
8 a
9
10 b
11 o
12 o
13 k
14 .
15
16 W
17 h
18 e
19 n
Reading Pointer
Writing Pointer
MAX

Fig. 4


Start
S1
Set Reading Pointer to Zero.
S2
RETURN Command ?
YES
NO
S3
Store Charact. Data in Buffer A.
S4
Increment Writing Pointer.
S5
Continue ?
NO
End
YES
S6
Writing Pointer>MAX ?
YES
NO

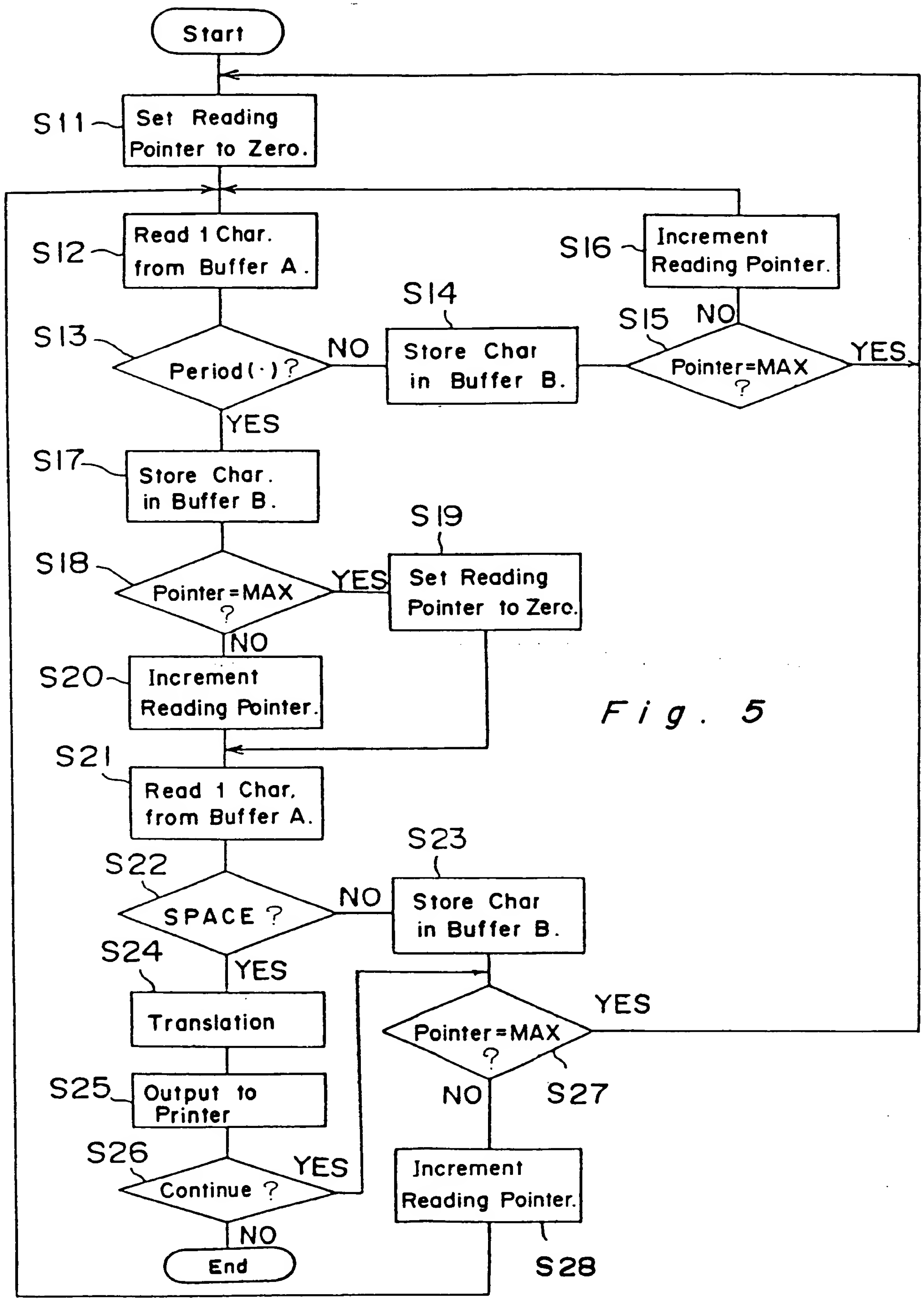

Start
S11
Set Reading Pointer to Zero.
S12
Read 1 Char. from Buffer A.
S13
Period(·)?
NO
S14
Store Char in Buffer B.
S15
Pointer=MAX ?
NO
YES
S16
Increment Reading Pointer.
YES
S17
Store Char. in Buffer B.
S18
Pointer=MAX ?
YES
S19
Set Reading Pointer to Zero.
NO
S20
Increment Reading Pointer.
S21
Read 1 Char. from Buffer A.
S22
SPACE ?
NO
S23
Store Char in Buffer B.
YES
S24
Translation
S25
Output to Printer
S26
Continue ?
YES
NO
End
Pointer=MAX ?
YES
S27
NO
Increment Reading Pointer.
S28

Fig. 5